# Supplementary Material

# The Local Resolution of Cryo-EM Density Maps

Alp Kucukelbir[1], Fred J. Sigworth[1,2], and Hemant D. Tagare[1,3]

[1]Department of Biomedical Engineering, Yale University, New Haven, United States

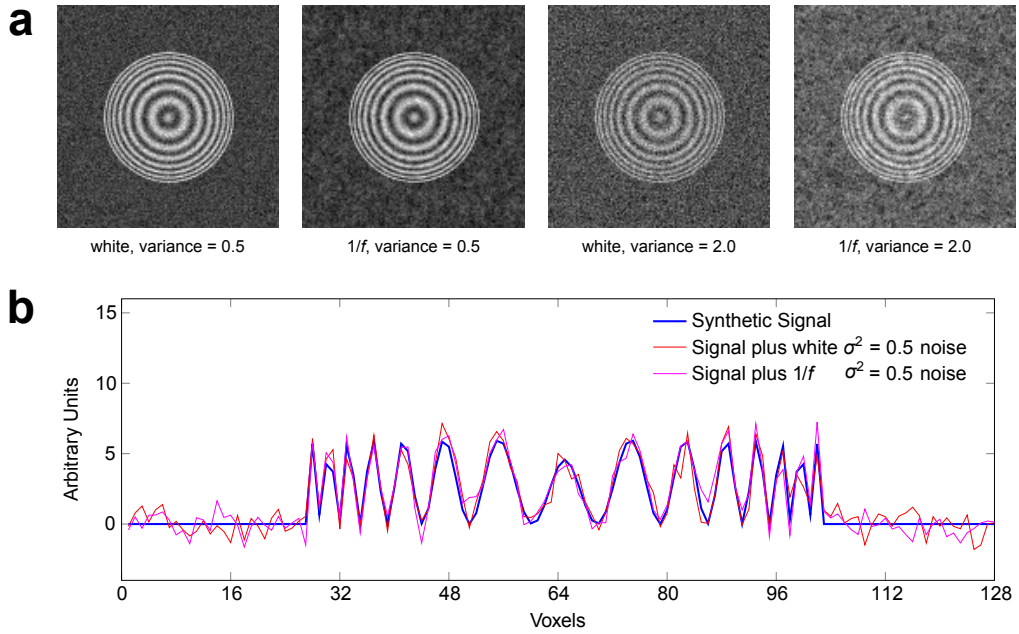
[2] Department of Cellular and Molecular Physiology, Yale University, New Haven, United States

[3] Department of Diagnostic Radiology, Yale University, New Haven, United States

Correspondence should be addressed to HDT (hemant.tagare@yale.edu)

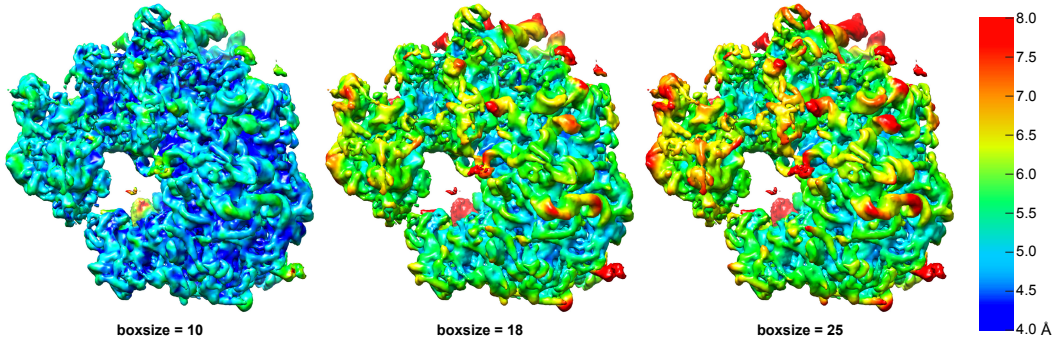
## Contents

# Supplementary Figure 1



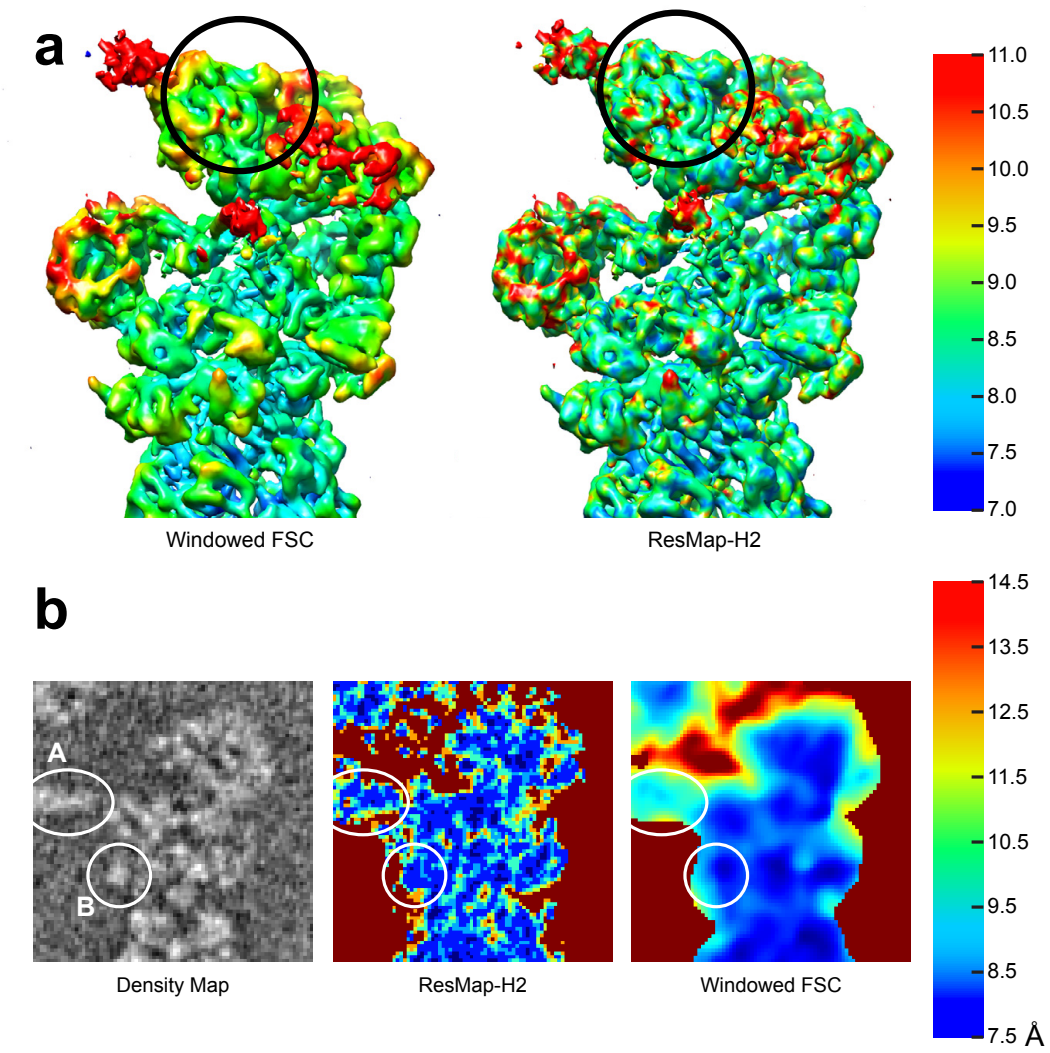

**Supplementary Figure 1** | Simulated data visualization. The simulated signal is $3 \times \cos(2\pi(1/(1-\sqrt[4]{R}+0.15))) + 3$ where $R$ is a radius map with zeros outside radius 38. We added both white and $1/f$ noise to the data. The purpose of the latter is to demonstrate the robustness of the local resolution estimates to non-white noise, which is typical of cryo-EM images[1]. **(a)** Central slices through simulated data corrupted by white and $1/f$ non-white noise at two variance levels. **(b)** Plot through central line for the variance level $\sigma^2 = 0.5$.

# Supplementary Figure 2



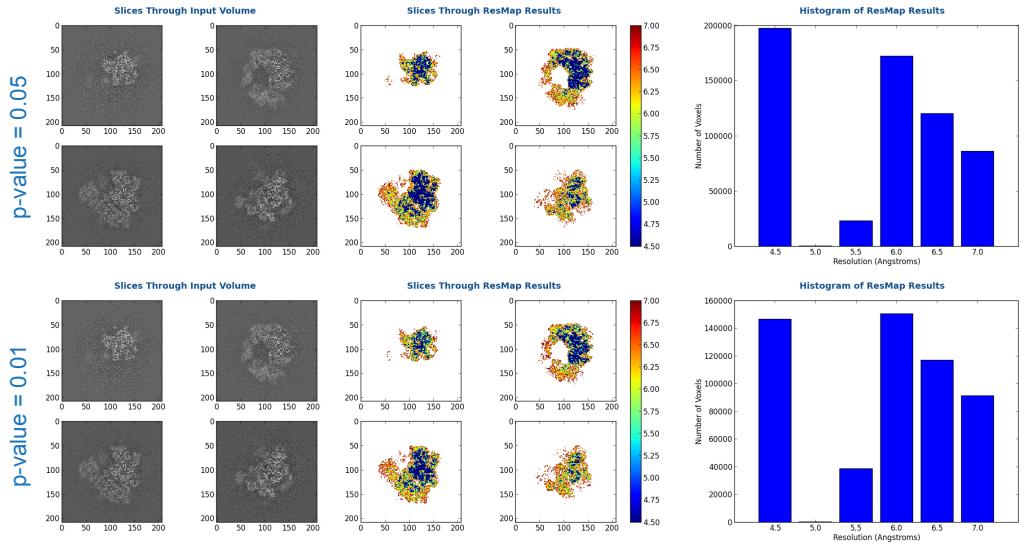

**Supplementary Figure 2** | Comparison of windowed FSC results. S.cerevisiae ribosome[2] (EMD-2275, split-dataset maps courtesy of the authors) volumes are rendered and colored with the output of blocres[3] using varying window sizes. Values are reported in Å. Results are sensitive to the boxsize parameter and take 25 minutes (boxsize = 10) to four hours (boxsize = 25) to compute on a four-core Fedora 14 workstation.

# Supplementary Figure 3



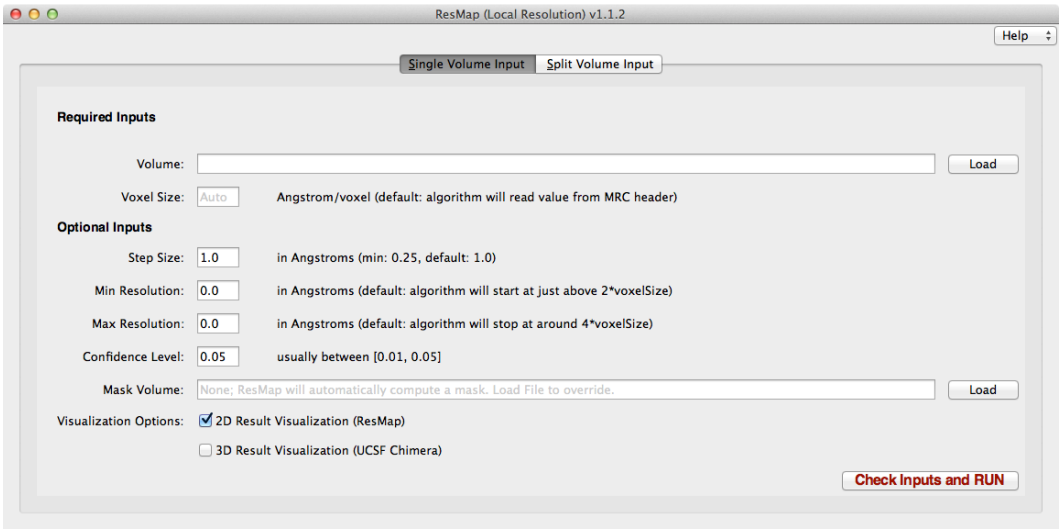

**Supplementary Figure 3** | Comparison of windowed FSC and ResMap-H2 results for the 26S Proteasome[4] (EMD-1992). Windowed FSC results were obtained using blocres[3] with boxsize = 18 (personal communication with Gabriel Lander). Values are reported in Å. **(a)** Volumes were median filtered in UCSF Chimera for visualization. Left: windowed FSC results. Right: ResMap-H2 results. The windowed FSC results estimate a lower resolution than ResMap-H2 in the area delineated in the black circle. This may be the result of windowed FSC using neighboring voxels in the solvent. **(b)** Central slice through unfiltered density map. Letters A and B highlight regions where the visual detail level appears to be comparable. ResMap-H2 corroborates this by assigning similar resolutions to both regions, while windowed FSC results appears to underestimate resolution in region A.

# Supplementary Figure 4




**Supplementary Figure 4** | ResMap behavior for typical statistical confidence levels. Results shown on the 80S ribosome[2] (EMD-2275). Colorbar values are reported in Å. As the p-value decreases, it is expected that fewer likelihood-ratio tests pass at smaller wavelengths. This is observed in the histograms above (note the different $y$-axis limits). Visualizations and histograms are generated from ResMap (**Supplementary Software**; http://resmap.sourceforge.net).

# Supplementary Figure 5




**Supplementary Figure 5** | Screenshot of ResMap's graphical user interface on Mac OSX. The volume to be analyzed is the only required input. The optional inputs may be used to analyze a specific resolution range and/or reduce computation time by taking larger step sizes. The algorithm will automatically delineate the particle from its surrounding solvent, though the option to override this using a 'mask volume' is also available. Further details are provided within the software package manual (**Supplementary Software**; http://resmap.sourceforge.net).

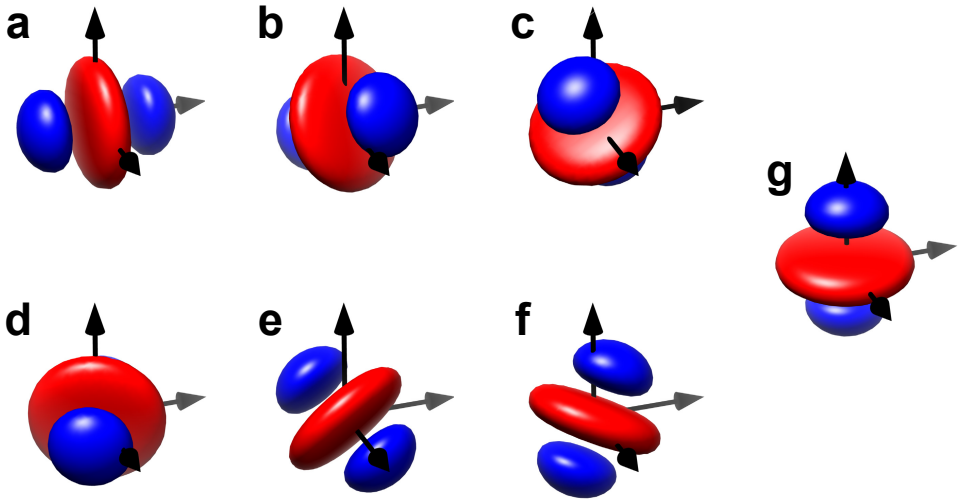
# Supplementary Note 1: Steerable Functions

Steerable functions are a finite set of functions with the property that linear combinations of these functions produce all possible rotations of themselves[5]. For example, the $G_{v,\alpha}$ function defined in the Online Methods section is a steerable function. In three dimensions, six $G^i_{v,\alpha}$'s oriented along the vertices of the icosahedron are sufficient to produce any arbitrary rotation of $G_{v,\alpha}$[6]. The icosahedrally oriented set of functions $G^i_{v,\alpha}$ for $i = 1, \cdots, 6$, are shown in **Supplementary Figure 6a-f**.

**Supplementary Figure 6g** shows a $G_{v,\alpha}$ function oriented along the $z$-axis. This function is not in the original set $G^1_{v,\alpha}, \cdots, G^6_{v,\alpha}$, but can be obtained as,

$$\begin{aligned} G_{v,\alpha} = \; & 0.6545 \times G^1_{v,\alpha} + 0.0955 \times G^2_{v,\alpha} - 0.25 \times G^3_{v,\alpha} + \\ & 0.0955 \times G^4_{v,\alpha} - 0.25 \times G^5_{v,\alpha} + 0.6545 \times G^6_{v,\alpha}. \end{aligned}$$

Similar results follow for $H_{v,\alpha}$ with 10 rotated versions, oriented along the faces of the icosahedron.




**Supplementary Figure 6:** $G^i_{v,\alpha}$ steerable functions. Red and blue indicate negative and positive parts, respectively. Subfigures **(a)** through **(f)** depict $G^i_{v,\alpha}$ for $i = 1, \cdots, 6$. Subfigure **(g)** shows an arbitrarily rotated $G_{v,\alpha}$ that can be obtained as a linear combination of the previous six functions.

## Supplementary Note 2: Scaling of the Basis Functions

For simplicity we consider the one-dimensional $G_{v,\alpha}$ function centered at $v_x = 0$. Extending the result to 2D and 3D is trivial and the centering does not affect the spectral peak. The one-dimensional $G_{v,\alpha}$ function is defined as

$$G_{v,\alpha}(x) = \left(4(\alpha x)^2 - 2\right) \exp\left(-\frac{(\alpha x)^2}{2}\right),$$

and its Fourier transform is

$$\mathcal{F}[G_{v,\alpha}](\omega) = -\frac{1}{\alpha}\left(4\left(\frac{\omega}{\alpha}\right)^2 - 2\right) \exp\left(-\frac{1}{2\alpha^2}\omega^2\right).$$

Taking the derivative of the Fourier transform and setting it to zero gives the maximum at

$$\omega = \pm\sqrt{\frac{5}{2}}\alpha.$$

To scale $G_{v,\alpha}$ such that its spectrum peaks at $\omega = 2\pi(1/\lambda)$ we set

$$\alpha = \frac{2\pi}{\lambda}\sqrt{\frac{2}{5}},$$

where $\lambda$ is the desired spatial resolution we seek in Angstroms.

We can use the same factor for the quadrature function $H_{v,\alpha}$, since it is a polynomial fit to the Hilbert transform of $G_{v,\alpha}$, which in this case is simply a phase shift [5].

## Supplementary Note 3: Likelihood Ratio Testing Framework

The model described in the Online Methods section implies the following likelihood of observing the density map $\mathbf{S}$ at voxel $v$ under a local-sinusoid model with scaling parameter $\alpha$,

$$\mathcal{L}(\mathbf{S} \mid \beta \,;\, v, \alpha) \propto \exp\left(-\frac{1}{2\sigma^2}\left\|\sqrt{\mathbf{W}^D_{v,\alpha}}(\mathbf{S} - \Phi_{v,\alpha}\beta)\right\|^2\right),$$

where $\propto$ means 'proportional to' and $\|\cdot\|^2$ is the sum of voxels squared.

The likelihood ratio is then simply the following ratio

$$\Lambda(\mathbf{S} \,;\, v, \alpha) = \frac{\max\{\mathcal{L}(\mathbf{S} \mid \beta \,;\, v, \alpha) : \text{ restrained to } \mathcal{H}_0\}}{\max\{\mathcal{L}(\mathbf{S} \mid \beta \,;\, v, \alpha) : \text{ restrained to } \mathcal{H}_1\}} = \frac{\mathcal{L}(\mathbf{S} \mid \widehat{\beta}_0 \,;\, v, \alpha)}{\mathcal{L}(\mathbf{S} \mid \widehat{\beta} \,;\, v, \alpha)}$$

where $\widehat{\beta}_0$ and $\widehat{\beta}$ are the coefficient vectors that minimize the weighted residual sum of squares under the local constant and local-sinusoid models, respectively.

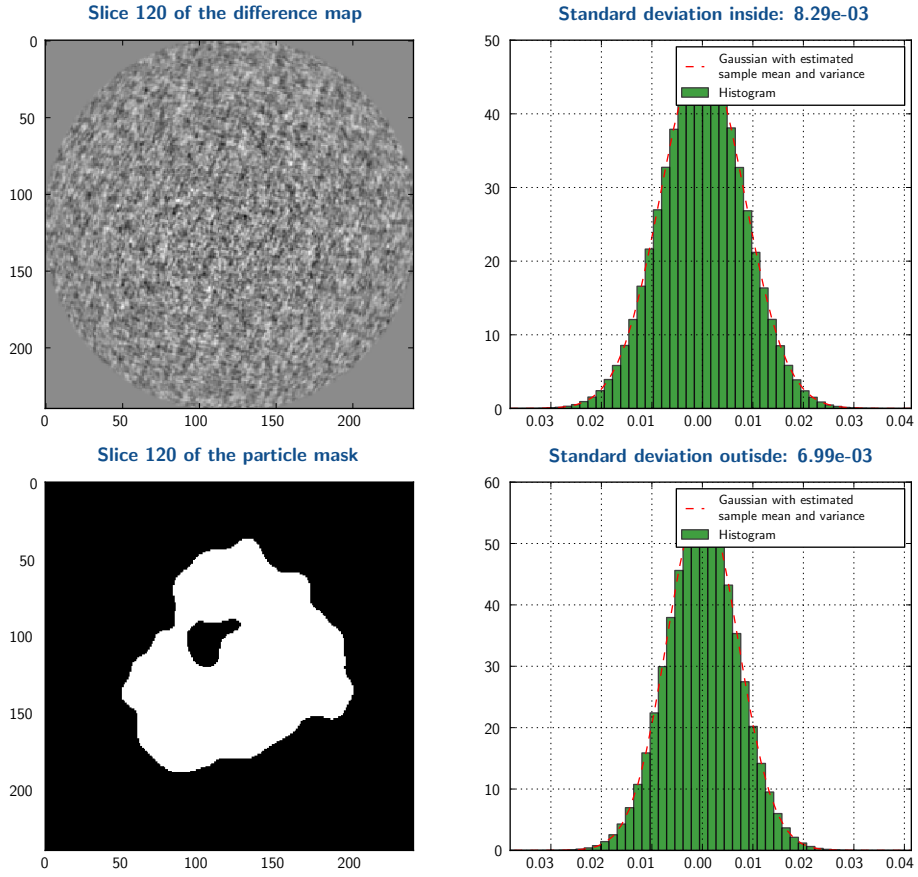
To ease computations, it is common to calculate the likelihood ratio statistic (LRS). Below we present the calculation of the LRS:

$$\begin{aligned}
\mathrm{LRS}(\mathbf{S} \,;\, v, \alpha) &= -2\ln\Lambda(\mathbf{S} \,;\, v, \alpha) \\
&= -2\ln\left(\mathcal{L}(\mathbf{S} \mid \widehat{\beta}_0 \,;\, v, \alpha)\right) + 2\ln\left(\mathcal{L}(\mathbf{S} \mid \widehat{\beta} \,;\, v, \alpha)\right) \\
&= \frac{1}{\sigma^2}\left(\left\|\sqrt{\mathbf{W}^D_{v,\alpha}}(\mathbf{S} - \mathbf{1}\widehat{\beta}_0)\right\|^2 - \left\|\sqrt{\mathbf{W}^D_{v,\alpha}}(\mathbf{S} - \Phi_{v,\alpha}\widehat{\beta})\right\|^2\right) \\
&= \frac{1}{\sigma^2}\left((\mathbf{S} - \mathbf{1}\widehat{\beta}_0)^\top \mathbf{W}^D_{v,\alpha}(\mathbf{S} - \mathbf{1}\widehat{\beta}_0) - (\mathbf{S} - \Phi_{v,\alpha}\widehat{\beta})^\top \mathbf{W}^D_{v,\alpha}(\mathbf{S} - \Phi_{v,\alpha}\widehat{\beta})\right) \\
&= \frac{\mathbf{S}^\top\Gamma_0\mathbf{S} - \mathbf{S}^\top\Gamma\mathbf{S}}{\sigma^2} \\
&= \frac{\mathbf{S}^\top(\Gamma_0 - \Gamma)\mathbf{S}}{\sigma^2} \\
\Gamma_0 &= \mathbf{W}^D_{v,\alpha} - \mathbf{W}^D_{v,\alpha}\mathbf{1}\left(\mathbf{1}^\top\mathbf{W}^D_{v,\alpha}\mathbf{1}\right)^{-1}\mathbf{1}^\top\mathbf{W}^D_{v,\alpha} \\
\Gamma &= \mathbf{W}^D_{v,\alpha} - \mathbf{W}^D_{v,\alpha}\Phi_{v,\alpha}\left(\Phi^\top_{v,\alpha}\mathbf{W}^D_{v,\alpha}\Phi_{v,\alpha}\right)^{-1}\Phi^\top_{v,\alpha}\mathbf{W}^D_{v,\alpha}
\end{aligned}$$

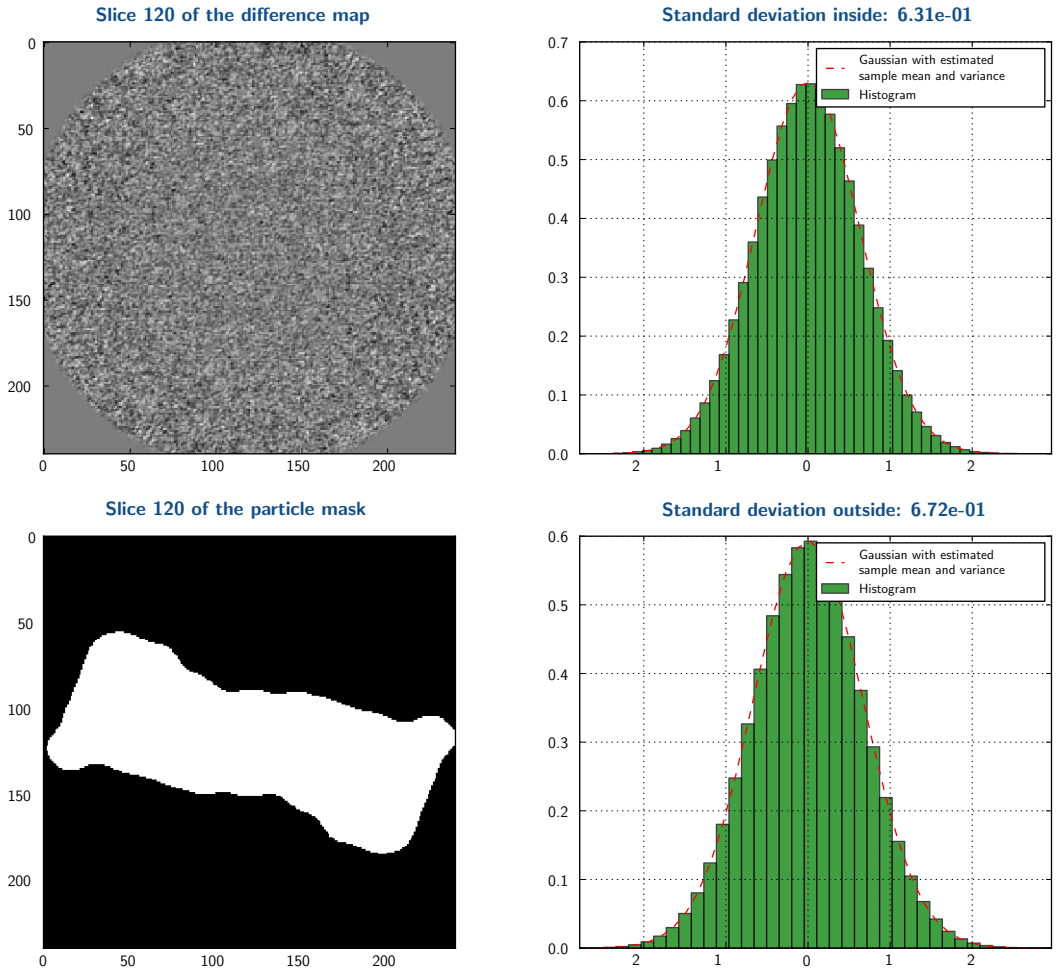
# Supplementary Note 4: Noise Statistics Inside and Outside the Particle

The theory presented in the Online Methods section is independent of the method used to estimate the noise variance. We use two noise estimators: one which analyzes the region outside the particle of a single reconstruction, and another which analyzes the difference between two split-dataset reconstructions in the region of the particle. Both approaches provide similar estimates of the noise variance and the noise statistics inside and outside the particle are nearly identical. The variance estimates agree to better than 20%, and their power spectra are indistinguishable for our purposes.

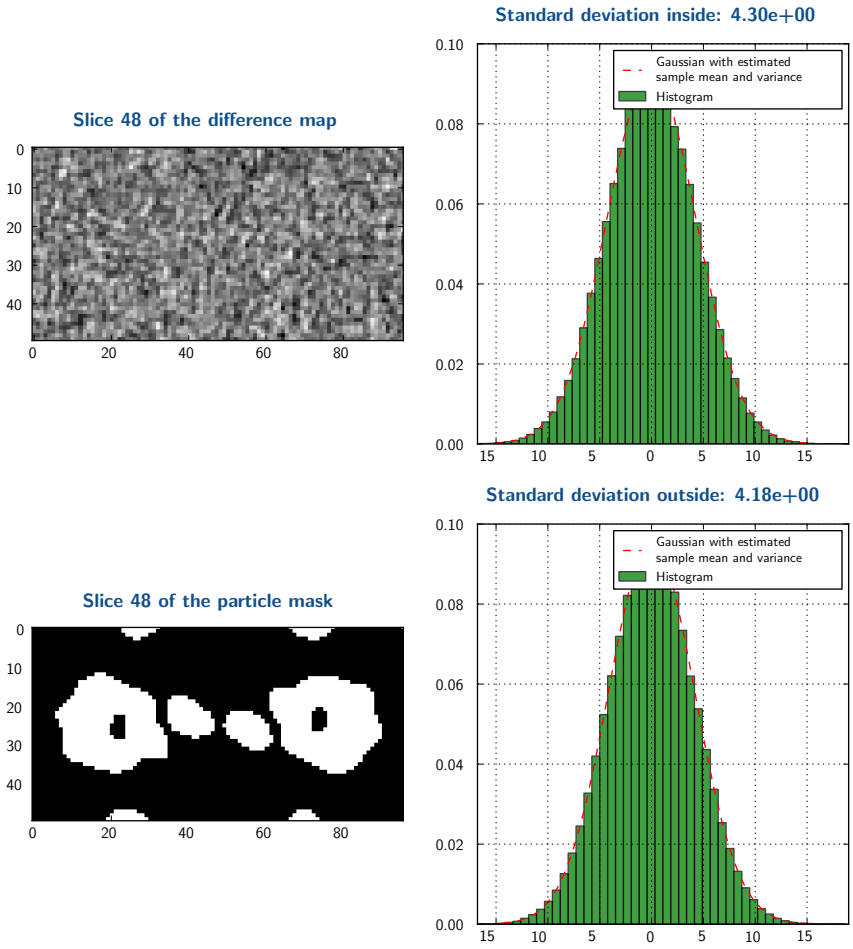
**Supplementary Figures 7-9** show the first order statistics of the noise inside and outside the particle region of three particles that span the 4 to 40Å range. **Supplementary Figures 10-12** show the corresponding noise power spectra.



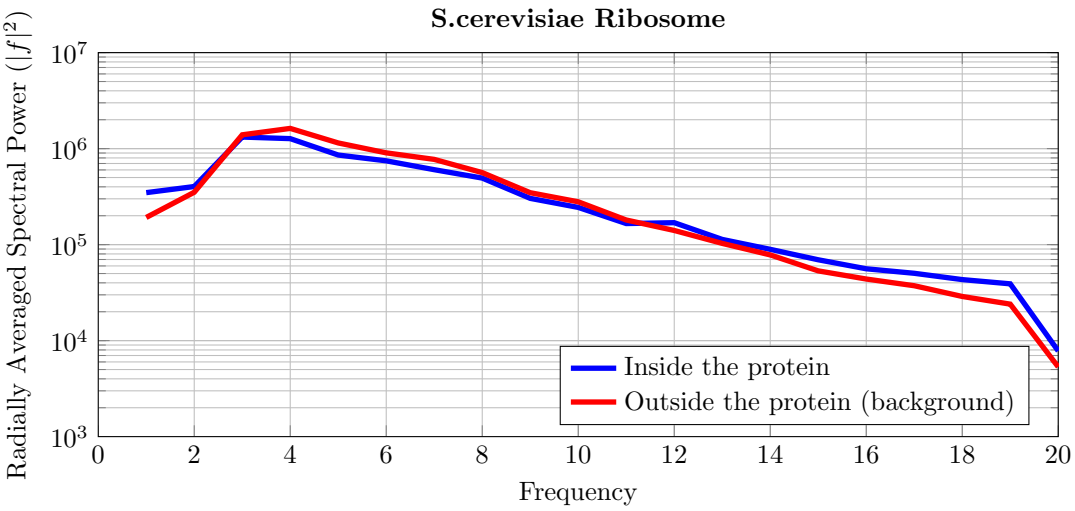

**Supplementary Figure 7:** Noise histograms for the S.cerevisiae Ribosome (EMD-2275). The difference map is computed by subtracting one of the split reconstruction from the other. The particle mask delineates where the protein is located. The histograms of the voxels show that the noise distribution follows a Gaussian shape, both inside and outside the particle. Moreover, the standard deviations of the histograms are very similar.



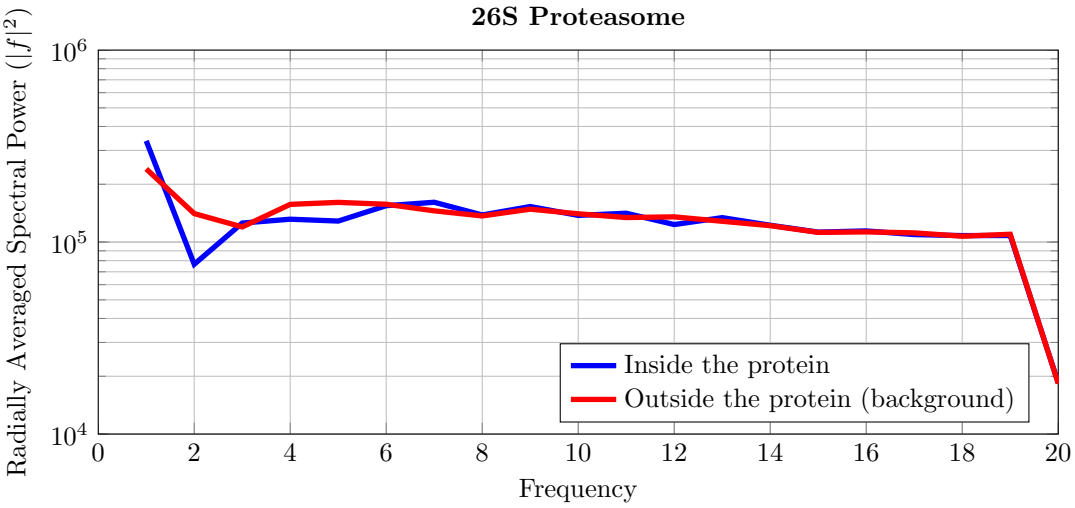

**Supplementary Figure 8:** Noise histograms for the 26S Proteasome (EMD-1992). The difference map is computed by subtracting one of the split reconstruction from the other. The particle mask delineates where the protein is located. The histograms of the voxels show that the noise distribution follows a Gaussian shape, both inside and outside the particle. Moreover, the standard deviations of the histograms are very similar.



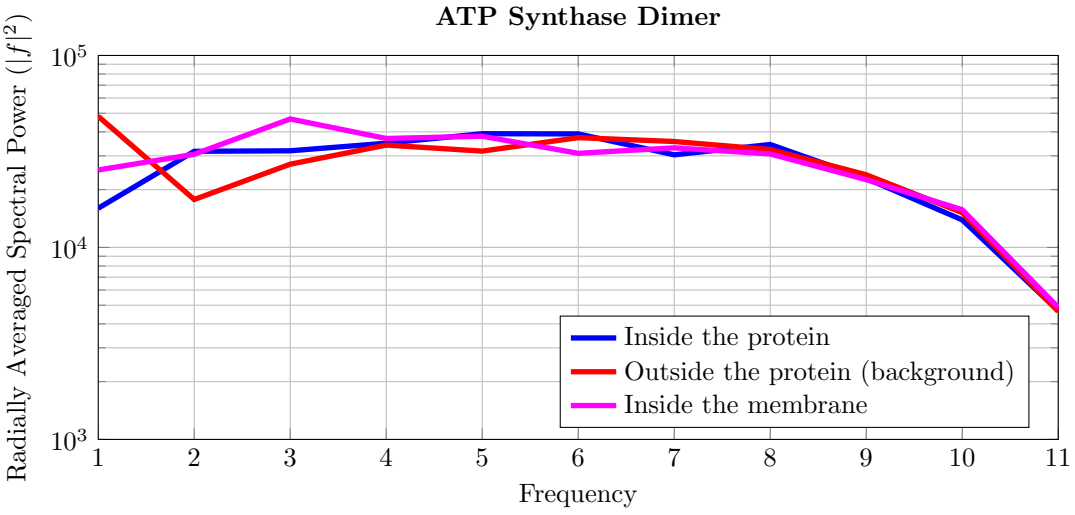

**Supplementary Figure 9:** Noise histograms for the ATP Synthase Dimer (EMD-2161). The difference map is computed by subtracting one of the split reconstruction from the other. The particle mask delineates where the protein is located. The histograms of the voxels show that the noise distribution follows a Gaussian shape, both inside and outside the particle. Moreover, the standard deviations of the histograms are very similar.




**Supplementary Figure 10:** Noise spectra for the S.cerevisiae Ribosome (EMD-2275). A box of size 40 voxels was placed inside and outside the protein mask and outside the protein. A 3D Fourier transform was computed for each box and then radially/spherically averaged to produce the plot above.




**Supplementary Figure 11:** Noise spectra for the 26S Proteasome (EMD-1992).A box of size 40 voxels was placed inside and outside the protein mask and outside the protein. A 3D Fourier transform was computed for each box and then radially/spherically averaged to produce the plot above.




**Supplementary Figure 12:** Noise spectra the ATP Synthase Dimer (EMD-2161). A box of size 22 voxels was placed inside the protein (i.e. where there is the ATP Synthase dimer), inside the membrane (i.e. where there is membrane density), and outside both. A 3D Fourier transform was computed for each box and then radially/spherically averaged to produce the plot above.

# References


[1] Baxter, W., Grassucci, R., Gao, H. & Frank, J. *J. Struct. Biol.* **166**, 126 (2009).

[2] Bai, X., Fernandez, I. S., McMullan, G. & Scheres, S. H. *eLife Sci.* **2** (2013).

[3] Heymann, J. B. & Belnap, D. M. *J. Struct. Biol.* **157**, 3 (2007).

[4] Lander, G. C. *et al. Nature* **482**, 186–191 (2012).

[5] Freeman, W. T. & Adelson, E. H. *IEEE Trans. Pattern Anal. Mach. Intell.* **13**, 891–906 (1991).

[6] Derpanis, K. G. & Gryn, J. M. *IEEE Inter. Conf. Image Process.* **3**, 553 (2005).